Supplemental material: Gaze Prediction using Machine Learning for Dynamic Stereo Manipulation in Games

George Alex Koulieris* Technical University of Crete INRIA

George Drettakis† INRIA

Douglas Cunningham‡ Technical University of Cottbus

Katerina Mania§ Technical University of Crete

1 Introduction

This supplemental document provides details on the various steps of our approach. Most of these are heavily based on standard methodology in the literature, but some adaptation was required to our specific problem.

2 Eye Tracking Data De-projection

The eye tracker yields time-coded [x,y] coordinates of fixations in the [0,1] range. To identify which object was fixated in the game FoV, we de-projected the eye-gaze space [x,y] coordinates in the frustum of the participant's dominant eye. A ray was then reconstructed originating from the dominant eye camera center and passing through the de-projected eye coordinates. The ray was advanced through the scene. The first non-transparent 3D model bounding volume that the ray hit, was considered the attended object.

3 Stereo Grading Details

We use the standard asymmetric viewing frusta, as presented among others [4] shown in Fig. 1.

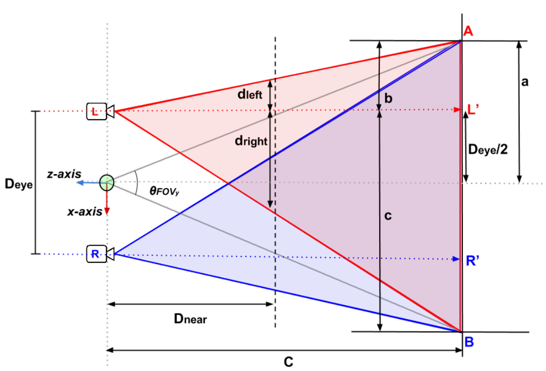

Figure 1: Asymmetric frustum stereo geometry.

In our system following the barycenter estimation of an object or a group of objects, the distance C is estimated as the barycenter + a user selected parameter that signifies how protruding or submerged the object/object will be perceived in relation to the zero parallax plane, i.e. the virtual screen. Distance D_{eye} is a user-defined seed value for our dynamic disparity control method.

To generate an asymmetric viewing frustum the near clipping plane's top, bottom, left and right coordinates in addition to the near and far clipping planes distances are required [3]. For the desired virtual screen a mono-frustum would be AOB. For this monoscopic frustum let us denote the Field of View (FOV) angle along the y-axis as θ_{FOVy} and the aspect ratio of the mono-frustum as r_{aspect}. We then estimate the parameters for both left and right frustums, in addition to D_{eye} as a system of simultaneous linear equations:

$$top = D_{near} \tan \frac{\theta_{FOVy}}{2} \qquad \& \qquad bottom = -top \tag{1}$$

The left frustum ALB, intersects the near clipping plane at d_{left} distance left of LL' and at d_{right} distance right of LL'. Given the triangles ALL' and BLL' we find that:

$$a = r_{aspect} C \tan \frac{\theta_{FOVy}}{2} \qquad \& \qquad \frac{d_{left}}{b} = \frac{d_{right}}{c} = \frac{D_{near}}{C} \tag{2}$$

$$b = a - \frac{D_{eye}}{2} \qquad \& \qquad c = a + \frac{D_{eye}}{2} \tag{3}$$

By interchanging b and c we estimate parameters for the right frustum ARB.

The image disparity p of a vertex with scene distance w is positive when the object is behind the virtual scene, and negative otherwise and is known as parallax. Parallax depends both on interaxial separation D_{eye} and convergence distance C. We estimate a comfortable D'_{eye} for a predetermined maximum on-screen parallax $|p|$ (see [2, 1]) based on user-display distance and display size:

$$D'_{eye} = \frac{w|p|}{w-c} \tag{4}$$

For the left eye frustum, parameters are:

$$left = -a + \frac{D'eye * D_{near}}{2C} \qquad right = a + \frac{D'eye * D_{near}}{2C} \tag{5}$$

For the right eye frustum, parameters are:

$$left = -a - \frac{D'eye * D_{near}}{2C} \qquad right = a - \frac{D'eye * D_{near}}{2C} \tag{6}$$

References

[1] G. R. Jones, D. Lee, N. S. Holliman, and D. Ezra. Controlling perceived depth in stereoscopic images. In *Photonics West 2001-Electronic Imaging*, pages 42–53. International Society for Optics and Photonics, 2001.

[2] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of vision*, 11(8):11, 2011.

[3] M. Woo, J. Neider, T. Davis, and D. Shreiner. *OpenGL programming guide: the official guide to learning OpenGL, version 1.2*. Addison-Wesley Longman Publishing Co., Inc., 1999.

[4] A. J. Woods, T. Docherty, and R. Koch. Image distortions in stereoscopic video systems. In *IS&T/SPIE's Symp. on Electronic Imaging: Science and Technology*, pages 36–48. International Society for Optics and Photonics, 1993.

*galekoul@gmail.com
†george.drettakis@inria.fr
‡cunningh@tu-cottbus.de
§k.mania@ced.tuc.gr